In vitro virus: Bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro

Naoto Nemoto[a], Etsuko Miyamoto-Sato[a,b], Yuzuru Husimi[c], Hiroshi Yanagawa[a,*]

[a] *Mitsubishi Kasei Institute of Life Sciences, Machida, Tokyo 194, Japan*
[b] *Yokohama National University, Hodogaya-ku, Yokohama 240, Japan*
[c] *Saitama University, Urawa, Saitama 338, Japan*

Received 30 June 1997; revised version received 1 August 1997

Abstract Adequate means for genotype assignment to phenotype is essential in evolutionary molecular engineering. In this study, construction of 'in vitro virus' was carried out in which a genotype molecule (mRNA) covalently binds to the phenotype molecule (protein) through puromycin on the ribosome in a cell-free translation system. Bonding efficiency was ~10%, thus indicating a population of the in vitro virus to have ~10^{12} protein variants, this number being 10^4 that in the phage display. The in vitro virus is useful for examining protein evolution in a test tube and the results may possibly serve as basis for a general method for selecting proteins possessing the most desirable functions.

Key words: In vitro virus; Genotype assignment to phenotype; Bonding of mRNA to protein; Puromycin; Evolutionary molecular engineering

1. Introduction

Genotype assignment to the phenotype is essential to Darwinian evolution. For this purpose in nature, the following modes of assignment are available: (i) ribozyme-type in which an identical molecule carries the genotype and phenotype, (ii) virus-type in which a genotype molecule is bound to phenotype molecules, as in the case of bacteriophage fd or M13; and (iii) cell-type in which genotype and phenotype molecules are present in identical compartments enclosed by a cell membrane.

'Evolutionary molecular engineering' using self-replicative RNAs was proposed in 1984 [1] and so far ribozyme-type assignment has proven appropriate for evolutionary molecular engineering owing to its simplicity [2]. But greater progress should be possible in this field if the genotype and phenotype are different molecules [1].

Bonding of a genotype molecule to the phenotype molecule is necessary for evolution to proceed smoothly. At present, the following has been proposed for evolutionary molecular engineering: phage display [3], polysome display [4], ribosome display [5], encoded combinatorial library [6], and cellstat [7]. But these do not permit global search in the sequence space of high dimension such as whole protein sequence space. For example, a virus such as bacteriophage can reproduce only in the host cells. This problem can be partially solved through use of a cell-free translation system.

In the present study, an 'in vitro virus' was constructed in which a genotype molecule (mRNA) binds to the phenotype molecule (protein) through puromycin in a cell-free translation system and will be amplified through RT-PCR. The life cycle of in vitro virus is a test tube version of that of retrovirus. The protein can be regarded as a coat protein of the virus. In vitro virus is a parasite to a cell-free test tube. It would evolve rapidly in it. Puromycin is a very powerful inhibitor of the growth of all cells, through its hindrance of chain elongation [8]. Its structure resembles the 3′ end of an aminoacyl-tRNA molecule, and thus is readily capable of entering the ribosomal A site to be transferred to nascent polypeptide chains by peptidyl transferase. Covalent bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of the encoded protein in a cell-free translation system using rabbit reticulocyte lysates is presented in the following.

2. Materials and methods

2.1. Chemicals and enzymes

The chemicals and enzymes were reagent grade and purchased from the following sources: puromycin from Sigma; rC-β amidite from PerSeptive Biosystems; Tetrazole from Nihon Milipore; tetrabutyl ammonium fluoride from Aldrich; T4 DNA ligase and T4 polynucleotide kinase from New England Biolabs; T4 RNA ligase and proteinase K from Takara; DNA polymerase (AmpliTaq Gold) from Roche; m^7G(5′)ppp(5′)G from BRL; T7 RNA polymerase, rabbit reticulocyte lysate translation system and mung bean nuclease from Promega; radiolabelled chemicals from Amersham. Other chemicals were of reagent grade.

2.2. Synthesis of rCpPur

rCpPur was produced by phosphoramidite chemistry [9] in the liquid phase. Puromycin (50 mg, 92 μmol) was made anhydrous by repeated coevaporation with dry pyridine (3×2 ml). This was followed by the addition of 15 ml of a solution of 4% tetrazole in acetonitrile and rC-β amidite (177.5 mg, 184 μmol). The reaction was monitored by silica gel TLC (solvent; chloroform:methanol = 9:1). After 1 day at room temperature, puromycin ceased to be apparent on TLC. The solvent was removed in vacuo. A 0.1 M solution (3 ml) of I_2 in tetrahydrofuran:pyridine:water (80:40:2) was added to oxidize the phosphite-triester and stirred at room temperature. After 1.5 h, the solvent was removed in vacuo and the residue extracted with chloroform. The extract was dried over $MgSO_4$ and the solvent removed in vacuo. The residue was chromatographed on a column of silica gel and eluted with chloroform containing 10% methanol to give the protected ribocytidyl-puromycin dimer (Rf 0.32 on silica gel TLC). For deprotection of all protecting groups, the protected dimer was first treated with 80% aqueous acetic acid (0.5 ml) for 1 h and the acetic acid was evaporated, was then added to 0.5 ml of a mixture of concentrated ammonia:ethanol (2:1). After 15 h ammonia and the

*Corresponding author. Fax: +81 (427) 24-6317.

Abbreviations: rC-β amidite, N^4-benzoyl-5′-*O*-(4,4′-dimethoxytrityl)-2′-*O*-*tert*-butyldimethylsilyl)-cytidine-3′-*O*-[*O*-(2-cyanoethyl)-*N*,*N*′-diisopropyl- phosphoramidite]; rCpPur, ribocytidyl-(3′-5′)-puromycin; TLC, thin layer chromatography; HPLC, high performance liquid chromatography; MALDI/TOF/MS, matrix-assisted laser desorption ionization/time of flight/mass spectrometry; P-acceptor, peptide acceptor

PII S0014-5793(97)01026-0

solvent were removed in vacuo and the residue was treated with 0.5 ml of 1 M tetrabutylammonium fluoride in tetrahydrofuran to remove the β-cyanoethyl group. After 30 min, the solvent was removed in vacuo and the residue chromatographed on QAE-Sephadex by a linear gradient of 0–0.5 M triethylamine carbonate. The eluent was collected and lyophilized twice. The dimer (10 mg) was obtained and analyzed by HPLC. Incubation of the dimer with nuclease P1 gave a digestion mixture of cytidine and puromycin 5′-phosphate in the correct ratios, which was analyzed by HPLC and then also by MALDI/TOF/MS (PerSeptive Biosystems Voyager). An $[M+H]^+$ molecular ion at m/z 777 for rCpPur was identified.

2.3. Construction of the genome of an in vitro virus

This genome is comprised of mRNA encoding a protein, DNA spacer, P-acceptor and rCpPur. The plasmid (pAR3040) containing the N-terminal half sequence of human tau protein [10] was used to construct the mRNA region. cDNA corresponding to amino acids 1–165 of a human tau protein was amplified by PCR using oligonucleotides primer 1 (5′ATGGCTGAGCCCCGCATGGAGTTC3′) and primer 2-1 (5′CTCTGCCACTTACTAGGGCTCCCG3′) with the stop codon or primer 2-2 (5′CTCTGCCACCTTCTTGGGCTCCC3′) without this stop codon. DNA sequences containing a T7 RNA polymerase promoter, Kozak sequence and sequence corresponding to positions 1–25 of a human tau protein were chemically synthesized. The purified DNA sequences were fused by PCR in two steps: DNAs first amplified in the absence of primers followed by amplification in the presence of primer 3 and primer 2 and purified fused DNA transcribed to mRNA by T7 RNA polymerase in vitro.

The P-acceptor (5′CTTACTGTCTTTTTTTTTTTTrGrArGrC3′), a 25 mer DNA-RNA chimera nucleotide having 4 mer RNA at the 3′-terminal end was ligated into a 105 mer DNA spacer (5′AAGCGACTCGCGTGGTCTCGCA (T×66) AATGACCTCCCC3′) by T4 DNA ligase in the presence of splint DNA at 16°C for 2 h. rCpPur was phosphorylated by T4 polynucleotide kinase and then ligated into the DNA spacer-P-acceptor by T4 RNA ligase at 37°C for 30 min. The DNA spacer-P-acceptor-rCpPur was phosphorylated by T4 polynucleotide kinase and ligated into the mRNA by T4 RNA ligase at 4°C for 48 h. Following extraction with phenol, the ligated product (in vitro virus genome) was precipitated with ethanol and then with polyacrylamide gel (4%)-purified at 60°C.

2.4. In vitro translation of the mRNA region of the in vitro virus genome

The mRNA region of the in vitro virus genome (0.3–1.2 μg) was translated for 20 min at 30°C in 50 μl containing 35 μl rabbit reticulocyte lysates [11]. Following translation, 1 mg/ml of vanadyl ribonucleoside complexes was added to the system followed by cooling on ice. The translation products were analyzed on 11.25% acrylamide-0.8% bisacrylamide-0.1% SDS gels which were exposed to an imaging plate for 30 min for analysis by an imaging analyzer (Fuji Film BAS2000).

3. Results and discussion

3.1. Construction of the genome of the in vitro virus

Fig. 1 shows the scheme for this construction that was conducted using mRNA encoding the N-terminal fragment (1–165) of human tau protein with or without a termination codon (ochre), DNA spacer, P-acceptor and rCpPur. Translation of the constructed genome gave polypeptides in a cell-free translation system. Puromycin of the 3′-terminal end of the genome entered the A site of the ribosome to covalently bind to the C-terminal end of polypeptide. It was then released from the ribosome and the in vitro virus particle (virion) in which the genotype was connected to the phenotype was obtained. This virion was found to be useful for selecting proteins possessing the most desirable functions.

The chain length between a termination codon and puromycin, presence of a termination codon and 3′-terminal structure of a genomic region were all taken into consideration for construction of the virus. These were also considered in the

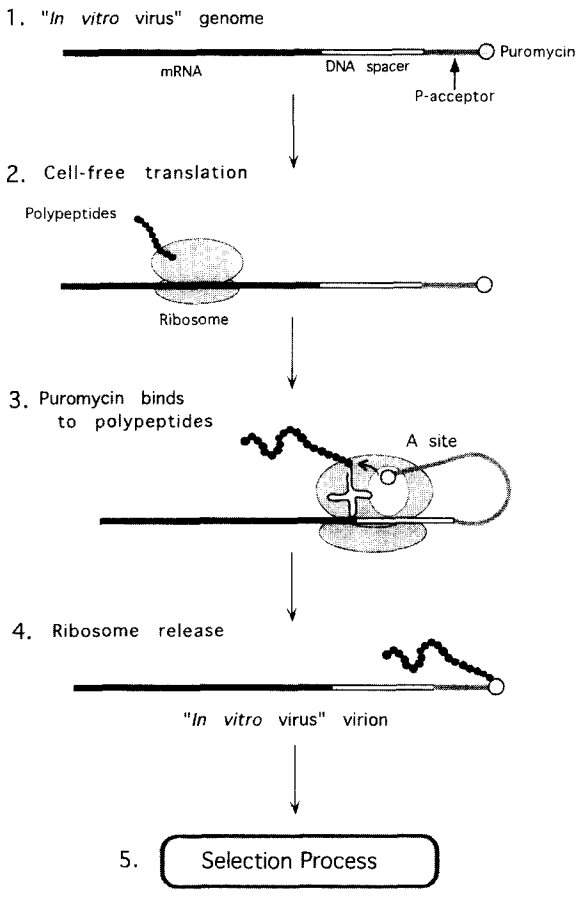

Fig. 1. Scheme for the construction of the in vitro virus virion.

subsequent construction of the following three different genomes of the in vitro virus: one with a termination codon and without a DNA spacer; one without a termination codon and DNA spacer; and finally, one without a termination codon and with DNA spacer. The intermolecular bonding of ^{32}P-labelled rCpPur to the C-terminal end of N-terminal fragment (1–165) of human tau protein was examined and the results are shown in Fig. 2. Regardless of the presence or absence of a termination codon on the genome without a DNA spacer, proteins with rCpPur attached to the C-terminal residue were synthesized to the same extent. The protein bonded to ^{32}P-labelled rCpPur coincided with a monomer of the protein incorporated ^{35}S-labelled methionine (the first right lane in Fig. 2). When a termination codon was absent and a DNA spacer present, the protein bonded to ^{32}P-labelled rCpPur was synthesized in considerable amount and three time as much was synthesized when the DNA spacer was not connected to the 3′-terminal end of mRNA. Ribosome pausing would thus appear to occur in DNA sequences, accompanied by rCpPur binding to the C-terminal end of the protein. These results indicate that a protein having rCpPur at the C-terminus is efficiently synthesized when a genome without a stop codon with DNA spacer was used as an mRNA in a cell-free translation system.

Termination is a relatively slow step involving a translational pause in rabbit reticulocyte lysates [12]. The rate of termination is one-tenth that of elongation [13]. Proteins translated from mRNAs without termination codon undergo modification by C-terminal addition of tag peptides each con-

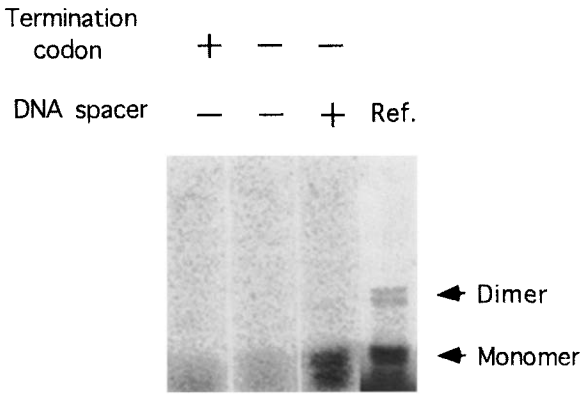

Fig. 2. Intermolecular bonding of ^{32}P-labelled rCpPur to the C-terminal end of N-terminal fragment (1–165) of human tau protein. Three in vitro virus genomes: one with a termination codon and without a DNA spacer, one without a termination codon and DNA spacer and one without a termination codon and with a DNA spacer, were translated in 50 μl containing 35 μl rabbit reticulocyte lysates at 30°C for 20 min. To the mixture 1.15 μg of the genome were added. The final concentration of ^{32}P-labelled rCpPur was 28 pM. Translation products were applied onto 11.25% SDS-PAGE. Ref (the first right lane), mRNA encoded the N-terminal fragment (1–165) of human tau protein was translated in the presence of [^{35}S] methionine under the same condition described above.

taining 11 amino acids [14]. 10Sa RNA capable of folding into a tRNA-like structure functions as mRNA and tRNA for tagpeptides. In this case 10Sa aminoacylated with alanine may be incorporated into the A site on ribosomes during ribosome pausing [15]. Thus, if ribosome pausing is relatively long, an tRNA-like molecule such as puromycin or 10Sa may possibly be incorporated into an apparently empty A site on the ribosome.

3.2. *Construction of an in vitro virus virion in a cell-free translation system*

This construction was carried out in a cell-free translation system. Examination was initially made of the effects of only mRNA encoding the N-terminal fragment (1–165) of tau protein on the incorporation of methionine in a cell-free translation system using rabbit reticulocyte lysates. As shown in Fig. 3A, ^{35}S-labelled methionine was incorporated into two proteins with molecular weights of ~28 kDa and ~55 kDa (the first left lane). The former corresponded to the apparent molecular weight of a monomer of the N-terminal fragment (1–165) of tau protein and the latter to that of its dimer. This indicates that mRNA functions as a messenger for the synthesis of N-terminal fragment (1–165) of tau protein in vitro. Study was then made of the effects of the genome, in which mRNA, DNA spacer (105 mer), P-acceptor and rCpPur are connected on the incorporation of methionine in the same translation system. Except for positions corresponding to the monomer and dimer, ^{35}S-labelled methionine was incorporated into a new broad band appearing at a position slightly higher than the in vitro virus genome (the first right lane). This broad band increased with incubation time (Fig. 3A) and amount of the genome (Fig. 3B). The genome of this type thus appears quite likely to bind covalently to the encoded protein through puromycin. For the genome with a DNA spacer, many protein bands appeared at regular intervals at the position of a monomer (Fig. 3A). Ribosome pause may thus possibly occur in the DNA spacer moiety of the genome followed by the pausing of many ribosomes to produce proteins with shorter chain lengths. Bands appeared at regular intervals near the monomer after 10 min incubation are not due to the cleaved mRNAs by contaminant RNases because the mRNA labeled with [^{32}P] at the 3′-terminus was not cleaved to give discrete bands appeared at regular intervals (lane 3 in Fig. 4).

The spacers in the genomes of in vitro virus had chains differing in length and the effects of this difference on bonding to encoded proteins were examined. Genomes with chain lengths of more than 100 mer bound to the encoded proteins but genomes with chain length less than 80 mer failed to do so (data not shown).

The covalent bonding of a genotype molecule (mRNA) to phenotype molecule (protein) was also studied using ^{32}P-labelled rCpPur in a cell-free translation system (Fig. 4). A genome was constructed, consisting of mRNA encoding the N-terminal fragment (1–165) of tau protein, DNA spacer (105 mer), P-acceptor and ^{32}P-labelled rCpPur. Covalent bonding of the genome having ^{32}P-labelled rCpPur at the 3′-terminal end to the C-terminal of the encoded protein was confirmed by treatment with mung bean nuclease after translation. Three bands were detected after this treatment (lane 4) and the two bands were assigned to the monomer and dimer of N-terminal fragment (1–165) of tau protein having ^{32}P-labelled rCpPur at the C-terminal end. The monomer and dimer detected after digestion with mung bean nuclase were completely digested with proteinase K and the top band in lane 4 was completely digested with RNase T_1 (data not shown). This result indi-

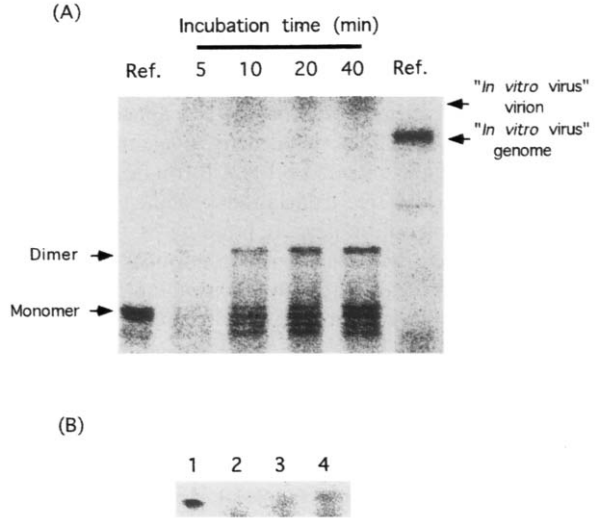

Fig. 3. Construction of in vitro virus virion in a cell-free translation system. A: Time course of intermolecular bonding of the in vitro virus genome consisting of mRNA, DNA spacer, P-acceptor and rCpPur to its encoded protein. The genome (0.48 μg) was translated in 50 μl containing 45 μCi [^{35}S] methionine and 35 μl of rabbit reticulocyte lysates at 30°C. Translation products were applied onto 11.25% SDS-PAGE. B: Effects of genome concentration on the intermolecular bonding of the genome to the encoded protein. Lane 1, genome with ^{32}P-labelled rCpPur at the 3′-end; lane 2, genome (1.2 μg) without rCpPur at the 3′-end; lane 3, genome (0.32 μg) with rCpPur at the 3′-end; lane 4, genome (0.64 μg) with rCpPur at the 3′-end. The genomes in lanes 2–4 were translated in 50 μl containing 45 μCi [^{35}S] methionine and 35 μl rabbit reticulocyte lysates at 30°C for 20 min but genome in lane 1 was not translated. The samples in lanes 1–4 were applied onto 11.25% SDS-PAGE containing 8 M urea.

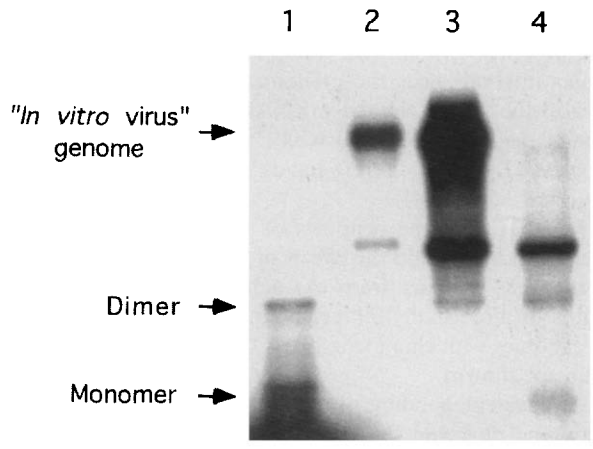

Fig. 4. Intermolecular bonding of the in vitro virus genome with ^{32}P-labelled rCpPur to the C-terminal end of the encoded protein. Lane 1, genome was translated at 30°C for 20 min in the presence of 45 μCi of [^{35}S] methionine; lane 2, genome with ^{32}P-labelled rCpPur at the 3′-end was not translated; lane 3, genome (1 μg) with ^{32}P-labelled rCpPur at the 3′-end was translated at 30°C for 20 min; lane 4, genome (1 μg) with ^{32}P-labelled rCpPur at the 3′-end was translated at 30°C for 20 min and then digested with mung bean nuclease (90 units) at 37°C for 10 min. Translation products were applied onto 11.25% SDS-PAGE.

cates that the monomer and dimer consist of a protein having rCpPur at the C-terminal end. The monomer band may possibly have been a contaminant in which the preformed monomer intermolecularly bonded to ^{32}P-labelled rCpPur remaining undergoing degradation from the genome in the mixture. But this was shown not to be the case since the monomer band did not appear before treatment with mung bean nuclease (lane 3). The formation of the dimer was stimulated by connection of a DNA spacer to mRNA (Fig. 3A). The in vitro virus genome consisting of mRNA, DNA spacer, P-acceptor and rCpPur may thus be concluded to covalently bind to the C-terminal end of the encoded protein through puromycin on the ribosome in a cell-free translation system.

The efficiency of intermolecular bonding of the genome to the encoded protein was ~10%. A solution of in vitro virus genome could be prepared at 40 to 100 pmol/ml, thus showing a population of the in vitro virus to contain $2.4 \sim 6 \times 10^{12}$ protein variants, this being 10^4 times the number in phage display [16]. Genotype assignment to the phenotype using the in vitro virus makes possible the synthesis of a great many mutants, production of different functional proteins, elimination of any permeability problem and the incorporation of different nonnatural amino acids. Meanwhile, in the study of the origin of life, a hypercycle with virus-like members was recently shown to make possible the emergence of the first encoded protein as a replicase cofactor and gradual evolution of translation activity out of the RNA world [17]. Thus the in vitro virus is useful for examining the initial protein evolution in the RNP world and for protein engineering in a test tube, and is an in vitro alternative to phage display.

Acknowledgements: The authors thank M. Nakamura for the synthesis of puromycin derivatives and T. Tsuji for the mass spectrometric measurement of puromycin derivatives.

References

[1] Eigen, M. and Gardiner, W. (1984) Pure Appl. Chem. 56, 967–978.
[2] Yarus, M. (1993) FASEB J. 7, 31–39.
[3] Smith, G.P. (1985) Science 228, 1315–1317.
[4] Mattheakis, L.C., Bhatt, R.R. and Dower, W.J. (1994) Proc. Natl. Acad. Sci. USA 91, 9022–9026.
[5] Hanes, J. and Plückthun, A. (1997) Proc. Natl. Acad. Sci. USA 94, 4937–4942.
[6] Brenner, S. and Lerner, R.A. (1992) Proc. Natl. Acad. Sci. USA 89, 5381–5383.
[7] Husimi, Y., Nishigaki, K., Kinoshita, Y. and Tanaka, T. (1982) Rev. Sci. Instrum. 53, 517–522.
[8] Traut, R.R. and Monro, R.E. (1964) J. Biol. Chem. 10, 63–72.
[9] Sinha, N.D., Biernat, J., McManus, J. and Köster, H. (1984) Nucleic Acids Res. 12, 4539–4557.
[10] Goedert, M., Spillantini, M.G., Potier, M., Ulrich, J. and Crowther, R.A. (1989) EMBO J. 8, 392–399.
[11] Pelham, H.R.B. and Jackson, R.J. (1976) Eur. J. Biochem. 67, 247–256.
[12] Wolin, S.L. and Walter, P. (1988) EMBO J. 7, 3559–3569.
[13] Lodish, H.F. and Jacobsen, M. (1972) J. Biol. Chem. 247, 3622–3629.
[14] Tu, G.-F., Reid, G.E., Zhang, J.-G., Moritz, R.L. and Simpson, R.J. (1995) J. Biol. Chem. 270, 9322–9326.
[15] Himeno, H., Sato, M., Tadaki, T., Fukushima, M., Uchida, C. and Muto, A. (1997) J. Mol. Biol. 268, 803–808.
[16] Scott, J.K. and Smith, G.P. (1990) Science 249, 386–390.
[17] Nemoto, N. and Husimi, Y. (1995) J. Theor. Biol. 176, 67–77.